Project 2 hotels bookings

Context

Have you ever wondered when the best time of year to book a hotel room is? Or the optimal length of stay in order to get the best daily rate? What if you wanted to predict whether or not a hotel was likely to receive a disproportionately high number of special requests?

This hotel booking dataset can help you explore those questions!

Content

This data set contains booking information for a city hotel and a resort hotel, and includes information such as when the booking was made, length of stay, the number of adults, children, and/or babies, and the number of available parking spaces, among other things.

All personally identifying information has been removed from the data.

Acknowledgements

The data is originally from the article **Hotel Booking Demand Datasets**, written by Nuno Antonio, Ana Almeida, and Luis Nunes for Data in Brief, Volume 22, February 2019.
The data was downloaded and cleaned by Thomas Mock and Antoine Bichat for #TidyTuesday during the week of February 11th, 2020.

Inspiration

This data set is ideal for anyone looking to practice their exploratory data analysis (EDA) or get started in building predictive models!

https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand

Project Tasks:

1- **EDA (Exploratory Data Analysis) (Phase 1):**
   a. Download and Read Data
   b. Change column names to remove whitespace and replace by (_) in one-line code.
   c. Make full report on quality of data , Missing , nulls , duplicated , top 3 freq values , outliers.
   d. Treat outliers if there are.

2- **Visualizations (Phase 2):**
   a. Univariate Analysis:

Univariate Analysis is a type of data visualization where we visualize only a single variable at a time. Univariate Analysis helps us to analyze the distribution of the variable present in the data so that we can perform further analysis

Visualizations, such as histograms, distributions, frequency tables, bar charts, pie charts, and boxplots,

b. Bivariate analysis:

is the simultaneous analysis of two variables. It explores the concept of the relationship between two variables whether there exists an association and the strength of this association or whether there are differences between two variables and the significance of these differences.

The main three types we will see here are:

1. Categorical v/s Numerical
2. Numerical V/s Numerical
3. Categorical V/s Categorical data

Visualizations, such as scatter, count plot,

c. Multivariate Analysis:

It is an extension of bivariate analysis which means it involves multiple variables at the same time to find correlation between them. Multivariate Analysis is a set of statistical model that examine patterns in multidimensional data by considering at once several data variable.

Visualizations, such as , scatter with hue , Heatmap

Requirements

1. For every type of analysis I want to visualize **5 different plots** From your choices which suits data.

2. search on method to save every plot and name it with visualize Title.

3- write comments and insight for every plot.

3- Send me (Notebook and 15 photos)

Notes:

Write comments for every code.

Grading from 20 marks .